Trinity: Early Experience, Successes, and Lessons Learned

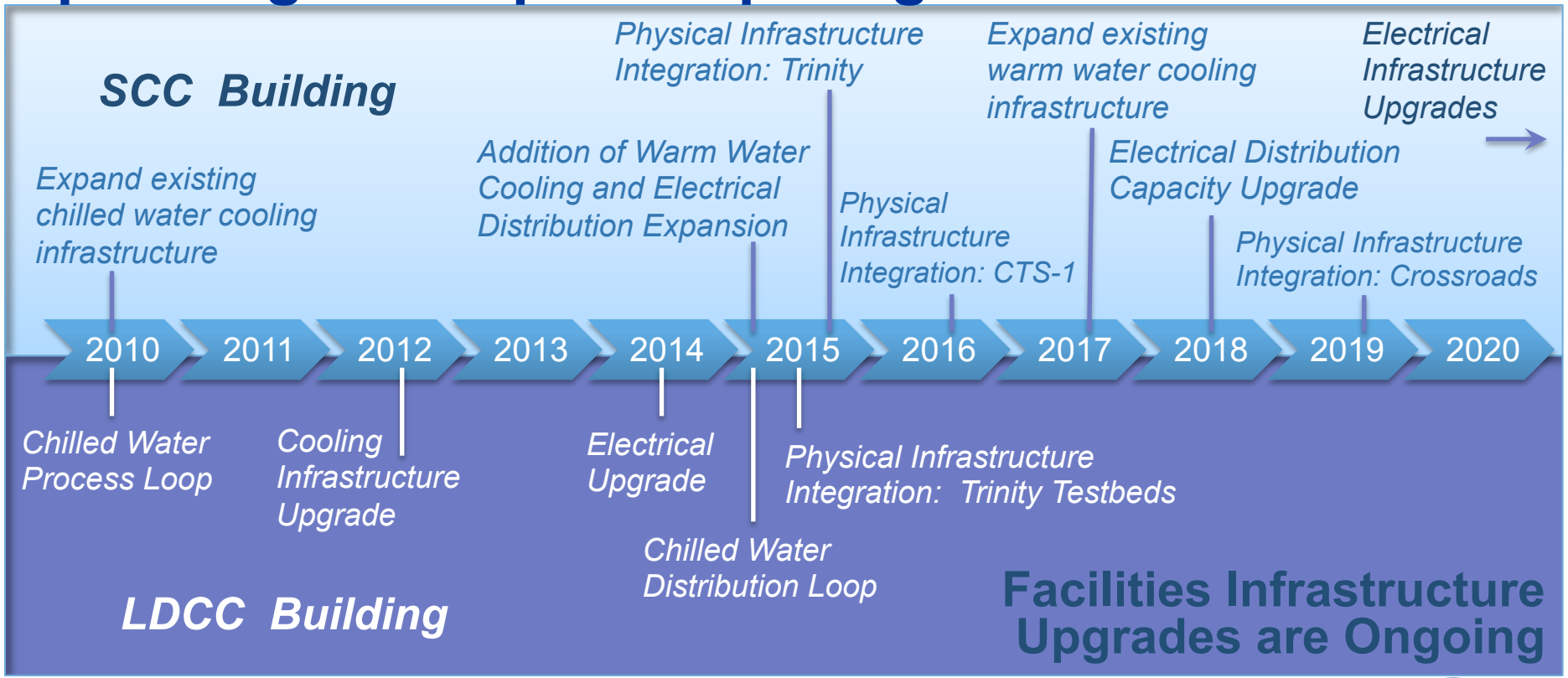
Carolyn Connor

HPC Facilities and Operations
Los Alamos National Laboratory
November 2015

Trinity: Early Experience, Successes, and Lessons Learned

This presentation gives a short summary of early experience, successes, and lessons learned with respect to facilities, operations, and monitoring of LANL's Trinity Supercomputer during the facility preparation and pre-acceptance testing phases of the project. Topics include facilities infrastructure upgrades, SERF water use, adaptive design and installation approaches, scalability and stability of monitoring system, and early power-capping investigation results.

HPC Supports LANL's Mission by Designing and Operating the Supercomputing Infrastructure

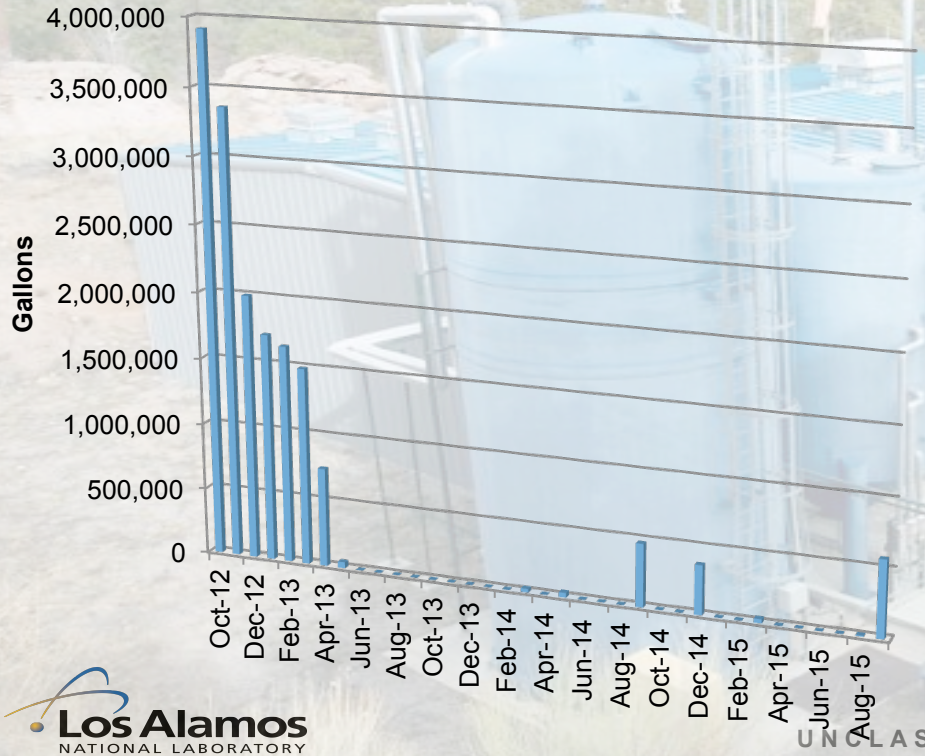


Currently the Sanitary Effluent Reclamation Facility (SERF) Supplies all Water to the SCC Data Center

The Sanitary Effluent Reclamation Facility treats effluent from the sanitary wastewater treatment plant.

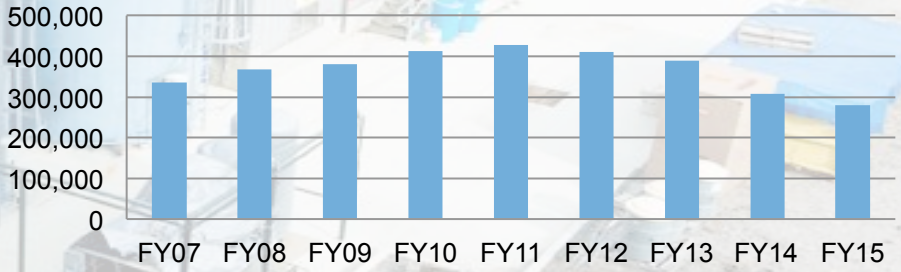
LANL Water Usage

SCC FY13-FY15 Potable Water Usage



- *FY15: record SERF water use (>30Mgal)*
- *SERF supplies all water needed for HPC in the SCC*

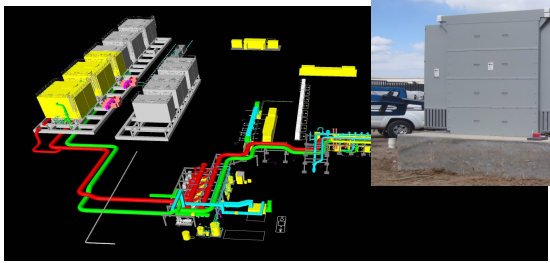
LANL Total Water Use (Kgal)



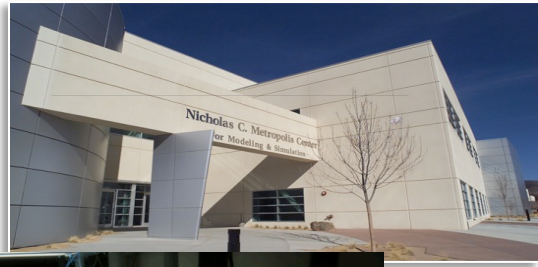
Adaptive Design and Installation Approaches Built on Historical Best Practices to Realize Efficiencies and Cost Savings

Power Distribution Approach:

- N+1 Available Power with Rotary UPS or Static UPS
  - File Systems and Disk Storage -- Network Switches
- Redundant Utility Power Available for Additional Reliability
  - Mechanical Systems -- Double-Ended Substations
- Raw Utility Power
  - Compute Racks
  - Building Automation System (BAS)

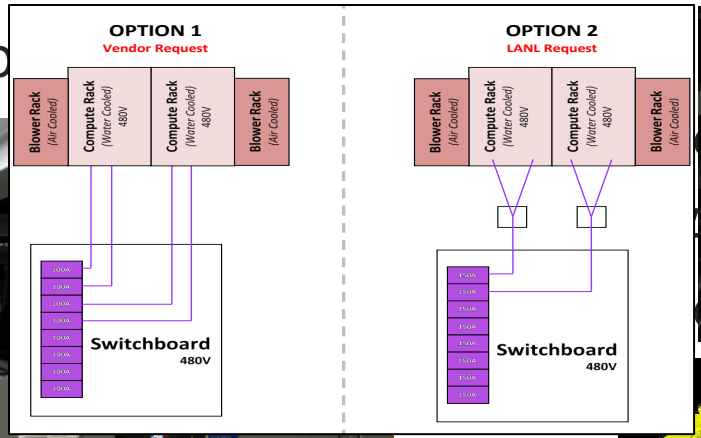


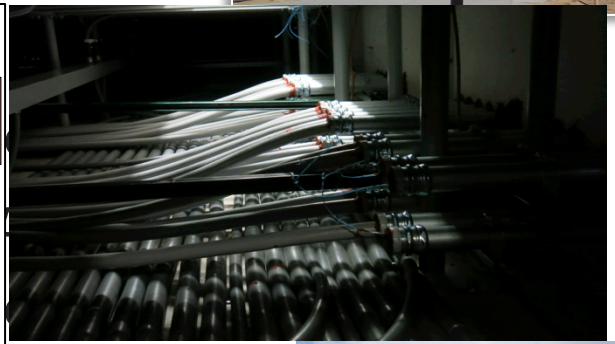
Adaptive Design and Installation Approaches Built on Historical Best Practices to Realize Efficiencies and Cost Savings



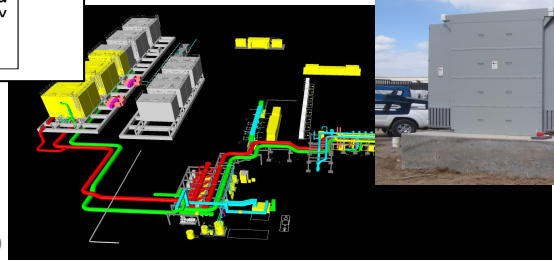
Power Distribution Approach:

- N+1 Availab





- Raw Utilit
  - Compu
  - Building m (BAS)

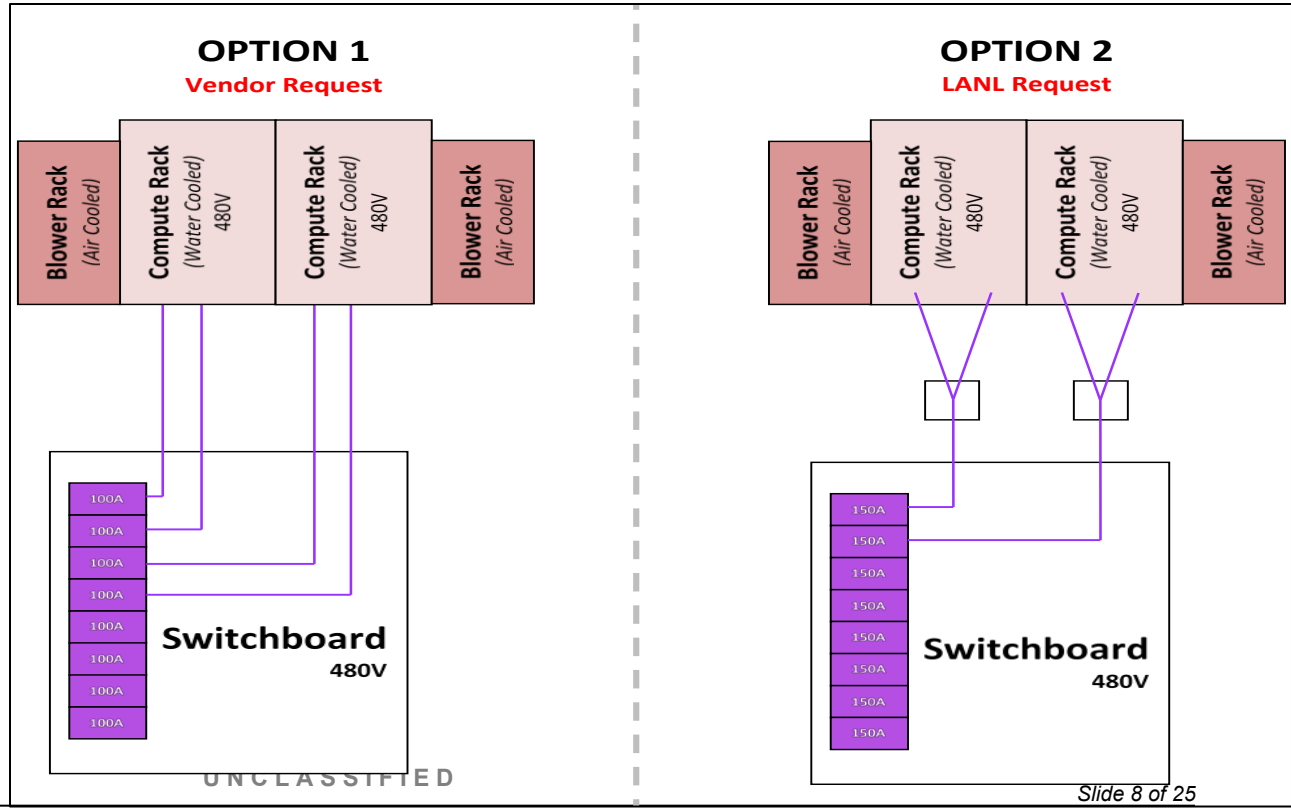


Adaptive Design and Installation Approaches: *J-Box Design Modification*

50% cost reduction

- material
- Labor
- Installation schedule

- Sandia adopted
- NERSC evaluating

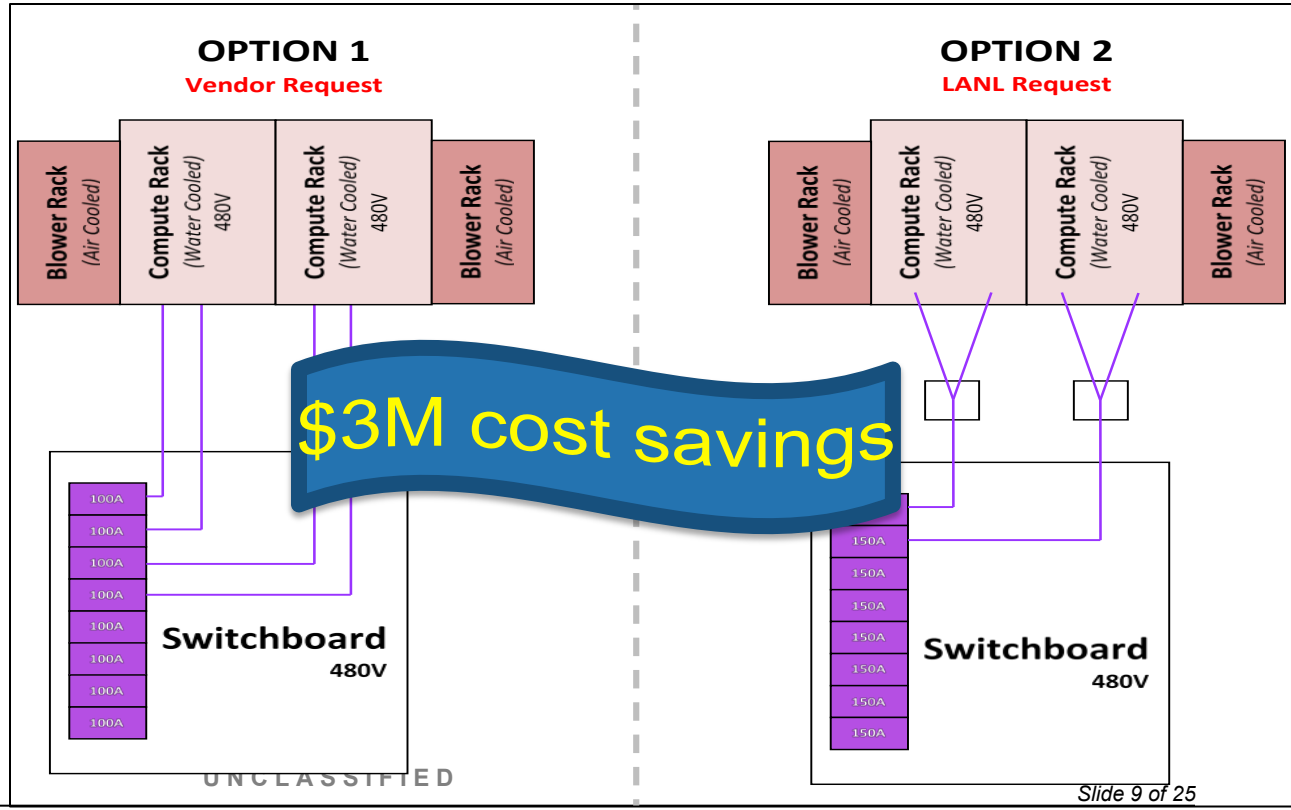


Adaptive Design and Installation Approaches: *J-Box Design Modification*

50% cost reduction

- material
- Labor
- Installation schedule

- Sandia adopted
- NERSC evaluating



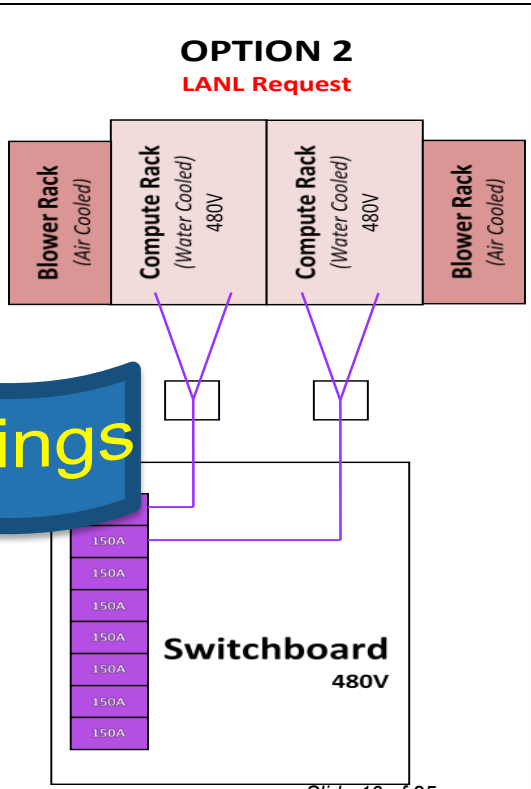
Adaptive Design and Installation Approaches: *J-Box Design Modification*

50% cost reduction

- material
- Labor
- Installation schedule

- Sandia adopted
- NERSC evaluating

$3M cost savings



Adaptive Design and Installation Approaches: *Smart Breakers*

- Installation of smart breakers enables rack-level data collection, including trending of power draw
- Validate integrity of power supplied to each rack
- Validate actual kW draw
- Improve issue tracking and isolation

Adaptive Design and Installation Approaches: Smart Breakers (tangible impact example: *Power Draw*)

Analysis of smart breaker data…

- Observed incidences of racks exceeding vendor's maximum power draw specification: 92 kW drawn during memory testing against 74 kW spec)
- Facility design implications
- Potential operational implications: tripped breakers, node shutdown, etc.

Adaptive Design and Installation Approaches: Smart Breakers (tangible impact example: *Power Distribution*)

Observation: recurring, unexplained incidences of machines powering down during acceptance testing. Why? → investigate the data

- Water (temp, flow, etc.) → nothing
- Electrical (smart breaker data)
  - Measured supply voltage to each rack
  - Cray: bench tested power supplies against LANL data → experimentally determined safe operating range (-25% to +5%)
  - LANL: Adjusted taps on transformers to guarantee operation within newly specified range, continuous monitoring
  - Cause: Power supply sensitivity on high end causing shutdown

Adaptive Design and Installation Approaches: Smart Breakers (tangible impact example: *Power Distribution*)

Observation: unexplained, recurring incidences of machines powering down during acceptance testing. Why? → investigate the data

- Water (temp, flow, etc.) → nothing
- Electrical (smart breaker data)
  - Measured supply voltage to each rack
  - Cray: bench tested power supplies against LANL data → experim… +5%)
  - LANL: … ion within newly specified range, continuous monitoring
  - Cause: Power supply sensitivity on high end causing shutdown

Vendor success story partnering

Solution (no further incidences)

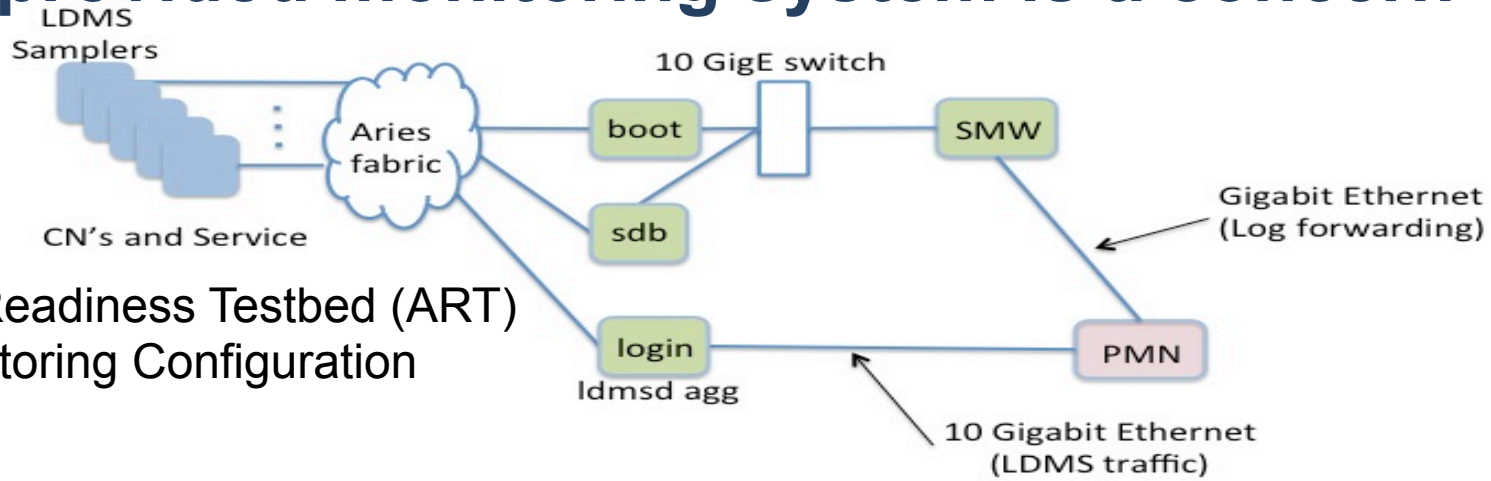
Adaptive Design and Installation Approaches:

TC Cable Preassembled alternative to in-house assembly of 24,000 linear feet of electrical cable

- IT Data Room Approved Cabling (worked through Fire Marshall and AHJ approval processes)
- Equivalent functionality and safety performance at significantly reduced cost
  - Installation time (1 month schedule gain)
  - Labor costs

$1M cost savings

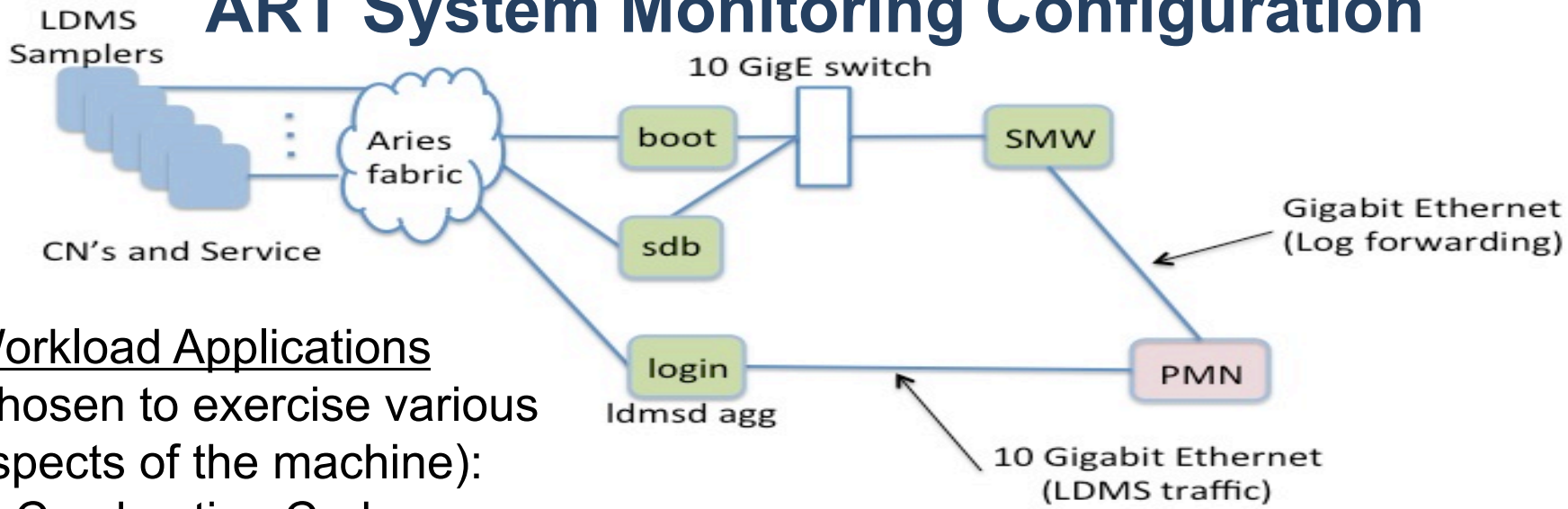
Lesson Learned: Stability and scalability of vendor-provided monitoring system is a concern



Application Readiness Testbed (ART)
System Monitoring Configuration

- No mechanism for checking heartbeat of monitoring network
- Vendor-provided monitoring network inadequate at Trinity scale
- ACES developing a mechanism to continuously remove data from SMW (bottleneck) for monitoring and analysis

ART System Monitoring Configuration



Workload Applications

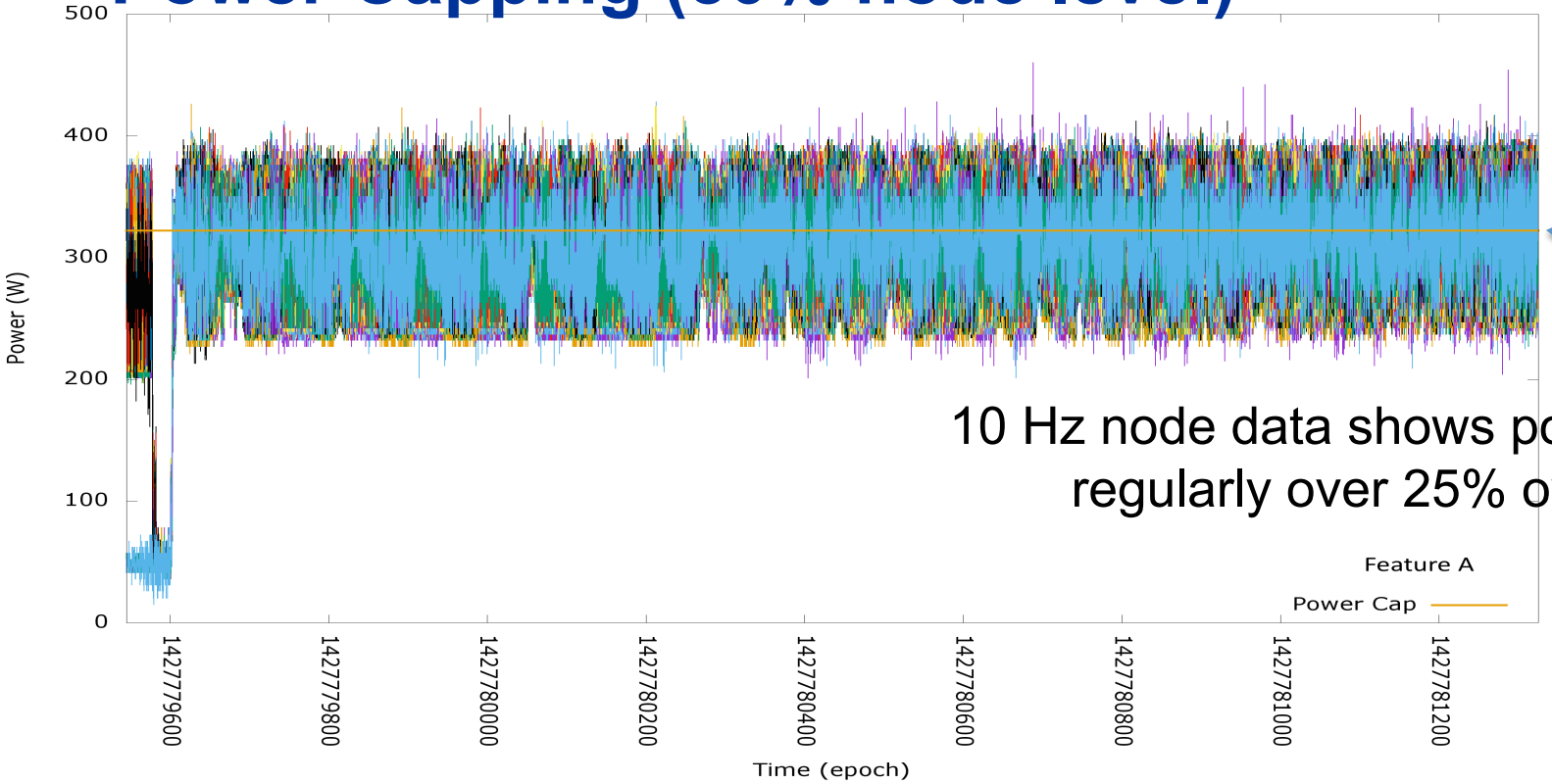
Chosen to exercise various aspects of the machine):

- Combustion Code – representative app
- HPCG – memory bound
- HPL - compute bound

Vary power capping focusing on:

- Power, cooling, temperature
- Facilities, machine, component data
- Behaviors, interactions, variations

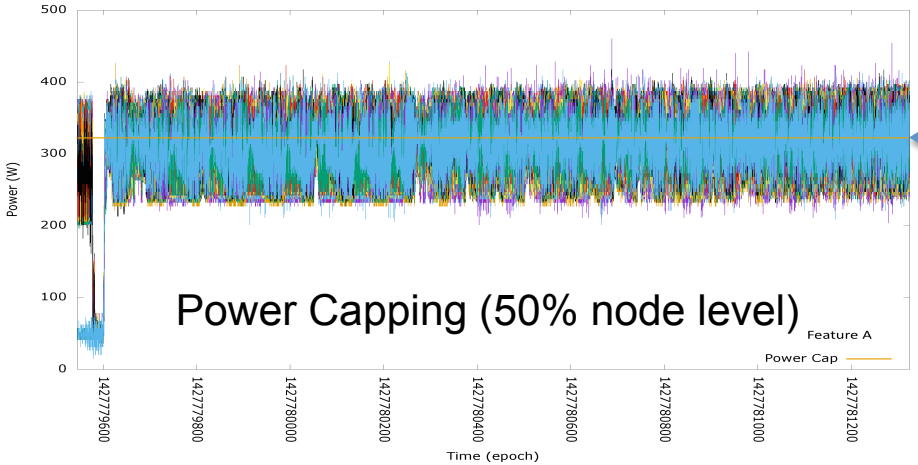
Power Capping (50% node level)

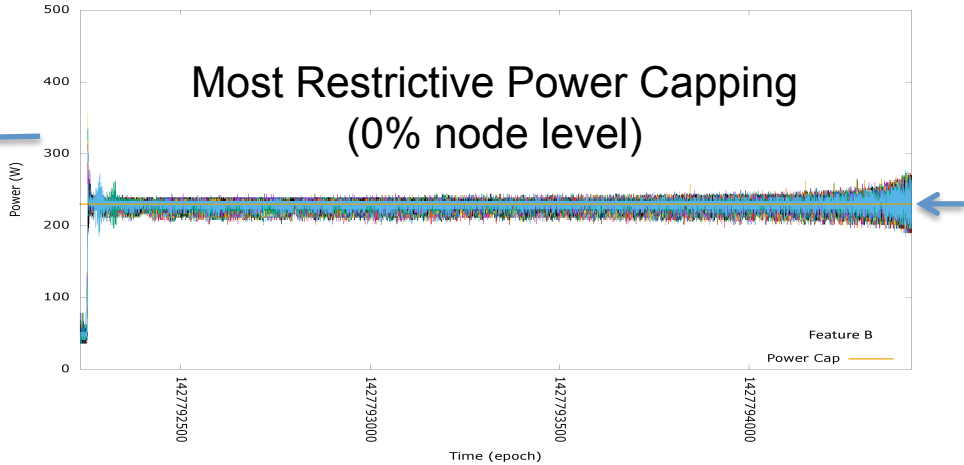


Power Cap

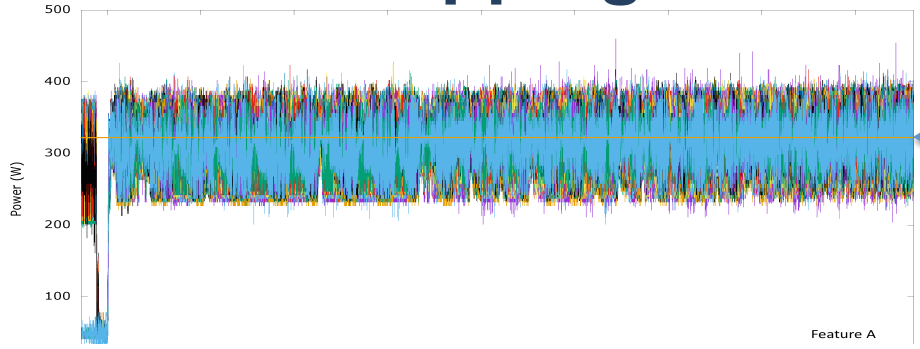
10 Hz node data shows power usage regularly over 25% over the cap

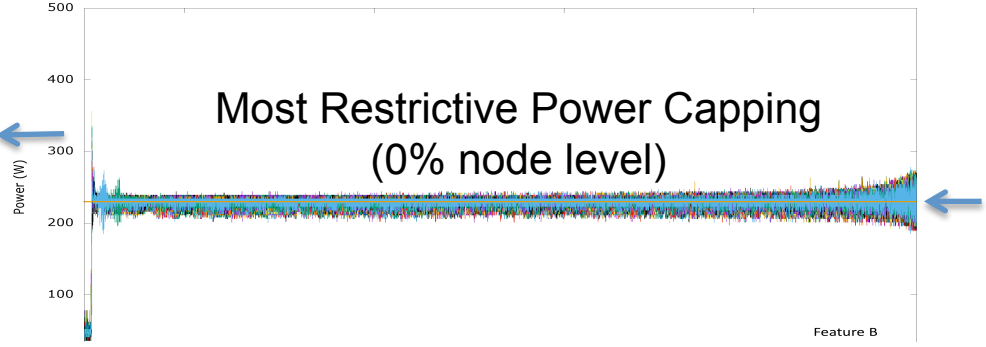
Power Capping

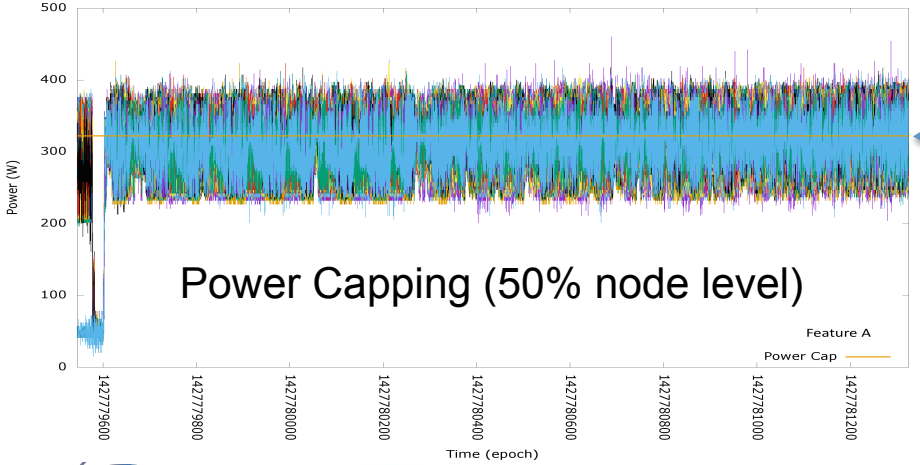


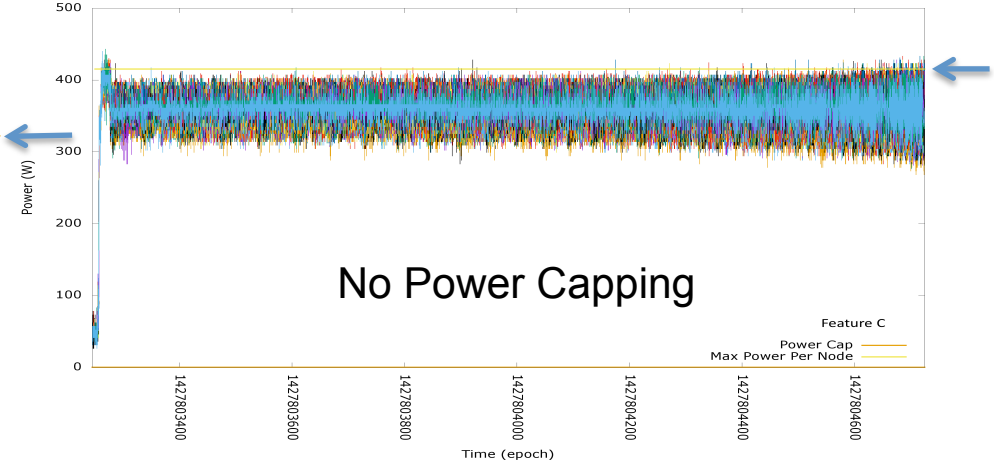


Power Capping

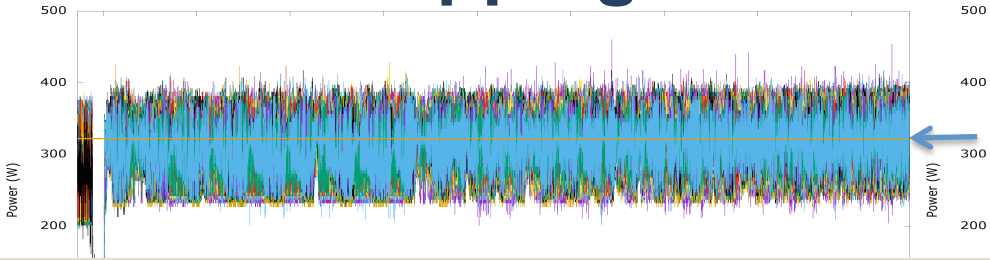


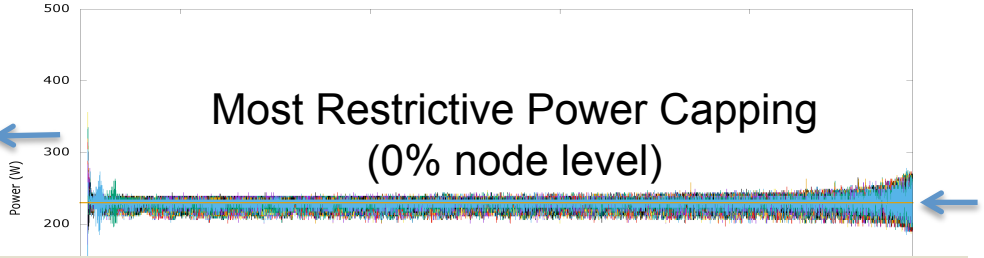






Power Capping





Conclusion: Power capping is not being pursued on Trinity at this time.

- Further investigation with Baler Log Analysis Tool corroborated experimental observations
- Anomalous, unexpected, and unexplained behavior that is not understood
- Interest in developing power capping and load shedding capability remains.
- Further analysis and understanding of system behaviors and abnormalities, particularly as they relate to power, thermal, and networking issues, is needed.

Challenges on the Horizon

- Working with Cray and Sandia to improve monitoring capabilities on Trinity
- Implementing the The Advanced Power Management System for Trinity
- Finding cost effective ways to implement data center efficiencies on older facilities

Acknowledgements

Thank you!

Teams (FOD, Facilities, CCS)

Cindy Martin, Ron Velarde, Jason Hunter, Phil Sena, Monica Witt, Andy Erickson, Eloy Romero, Michael Ferguson, Alynna Wiuff-Montoya, Dan O'Rourke, Tom Duran, Leonard Maez, Amanda Bonnie, Curtis Storlie, Sarah Michalak, Bill Rust, Anthony Lopez, Andrew Montoya and many more.

Study Results: Enabling Advanced Operational Analysis Through Multi-Subsystem Data Integration on Trinity (CUG 2015)

SNL: J. Brandt, D. DeBonis, A. Gentile, D. Martinez, S. Olivier, K Pedretti; LANL: C. Martin, J. Lujan, R. Velarde; Open Grid Computing: N Taerat

Sponsors: ASC, LANL Institutional Computing